IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:     Frudakis and Shriver          Art Unit:  1631
Application No.:  10/644,594               Examiner: Whaley, Pablo S.
Filed:         August 19, 2003             Conf. No.  6207
Title:         COMPOSITIONS AND METHODS FOR INFERRING ANCESTRY

Mail Stop: AMENDMENT
Commissioner for Patents
Washington. D.C.  20231

## DECLARATION UNDER 37 C.F.R. § 1.132

I, Tony Frudakis, Ph.D., declare and state that:

1.      I am a co-inventor of the subject matter described and claimed in U.S. Patent Application Serial No. 10/644,594, filed August 19, 2003, entitled "COMPOSITIONS AND METHODS FOR INFERRING ANCESTRY."

2.      I am familiar with the prosecution history of Patent Application Serial No. 10/644,594.

3.      I understand that the Examiner has rejected claims 1, 84-86, 90, 92-97, 100, 104, and 105 under 35 U.S.C. §103 (a), as allegedly being obvious over Parra et al. (Am J Physical Anthropol (2001) 114(1):18-29), in view of Ott et al. (Human Mutation (2001) 17:285-288) and Halushka et al. (Nature Genetics (1999) 22:239-247); claims 87-89 and 110-115 under 35 U.S.C. §103 (a), as allegedly being obvious over Para et al. (2001), in view of Ott et al. (2001) and Halushka et al. (1999), and in further view of Sorenson et al. (US Pub. No.: 2003/0172065, filed March 29, 2002);  claims 97-99, 101-103, and 106-109 under 35 U.S.C. §103 (a), as allegedly being obvious over Parra et al. (2001), in view of  Ott et al. (2001) and Halushka et al. (1999), and in

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page 2

PATENT
Attorney Docket No.: DNA1170-2

further view of Kanetsky et al. (Am J Hum Genet (2002) 70:770-775), Pritchard et al. (Theoretical Population Biology (2001) 60:227-237), and Pritchard et al. (Genetics (2000) 155:945-959); claims 1, 83-91, 95-99, and 107-110 under 35 U.S.C. §103 (a), as allegedly being obvious over Shriver et al. (Am J Hum Gent (1997) 60:957-964), in view of Daly et al. (Nature Genetics (2001) 29:229-232) and Kruglyak (Nature Genetics (1997) 17:21-24).

4.     The Examiner states repeatedly that previous works have made the present method for inferring genomic ancestral proportions and admixture obvious. Yet in none of these papers is ancestry being measured. The study of gene or phenotype histories is a very different thing from the study of ancestral histories because each gene has a unique history, subject to differing gene-specific environmental constraints. These environmental constraints are only correlated with (not deterministic of) ancestry – and often, the correlation is weak. In contrast, the coalescence of human populations -their relationships to one another and chronology of origins - is fixed in history, and indelibly recorded in the extra-genic regions of our genome.

The unequal distribution of gene sequences creates a different form of population structure than that due strictly to ancestry, and it is the latter that Applicants claim to be the first to measure. When the prior art uses markers found within gene regions as "Ancestry Informative" markers, relationships among human populations and individuals are measured that are a function of shared environments, convergent or coincident evolution. These measures are what we call "confounded ancestry" – not really ancestry per se, but relationships based on a mix of phenotypes, geography and environment.

Up until recently, appreciation of the distinction between "confounded ancestry" and "genomic ancestry" was far from obvious. The prior art recognized the difference, but the pervasive opinion at the time was that the latter simply did not exist for human populations (see, e.g., Parra et al., Proc Natl Acad Sci U S A (2003) 100(1):177-82). The prior art argued that "race" had no biological basis and that distinctions among human populations were the result of differing environments and diversity/adaptation of phenotypes rather than bona-fide population structure (see, e.g., Cho and Sankar, Nature Genetics (2004) 36:S8-S12). That the present

GT\6555018.1
331832-21

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page 3

PATENT
Attorney Docket No.: DNA1170-2

invention is novel an unobvious, and therefore relevant to this old debate, is attested to by the fact that Applicants, as leaders and pioneers in the field, were invited by the journal Nature Genetics to respond to the Cho and Sankar article of 2004, protesting the existence of ancestral population structure in the human population (see, Shriver et al., Nature Genetics (2005) 37:449-450).

The distinction between confounded ancestry and genomic ancestry is not straightforward, and required significant work. Proof that one is measuring the latter rather than the former requires a number of observations from highly engineered marker panels comprised of neutral markers, found outside of gene regions (not merely "non-coding"). Up until the present application, and immediately preceding completion of the first human genome draft, such assemblies were just not feasible and none of the prior art cited by the Examiner prior to 2002 meet the criteria of "neutral" (nor do they claim to). The distinction also requires the application of these markers in large-scale population surveys, which are difficult to carry out. The scientific community needed to observe that measures of "ancestry" were not strictly coupled with phenotype or geography, even though they were correlated with both as well as with "race." Population surveys on a global scale are necessary to demonstrate this and the inventors were the first to conduct such population surveys using bona fide AIMs (neutral SNPs; Frudakis, *Molecular Photofitting: Predicting ancestry and phenotype using DNA*, 2007 Elsiever/Academic Press Publishers, Burlington, MA). For example, if population structure due to genomic ancestry exists (as opposed phenotype or environment) scientists needed to be able to resolve groups such as South Asians and African populations, which the linguistics and paleoarchealogical record suggests should be possible, notwithstanding their similarly high levels of eumelanin in the epidermis (dark skin). To prove that bona-fide ancestry, rather than shared environmental history was being measured, it needed to be demonstrated that, unlike the inheritance of most anthropometric phenotypes, the inference of genomic ancestry in offspring from a knowledge of that of the parents was predictable, stable and reliable in a quantitative sense as Applicants expected. Applicants expected the inheritance of genomic ancestry to be more readily predictive because ancestry information is written on vast regions of each chromosome of an individual,

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page  4

PATENT
Attorney Docket No.: DNA1170-2

whereas phenotypes are caused by small sets of gene combinations, subject to higher sampling effects among gametes. Demonstration of predictability achieved by Applicants required detailed study in family pedigrees. It needed to be demonstrated that there existed variation in phenotype (as well as geography) among individuals of the same ancestry proportions. This is because ancestry is only correlated with phenotype, whereas causal gene sequences are linked to phenotype. Achieving the statistical power to do these things requires measures of "true" or genomic ancestry at the level of the individual – using populations of individual estimates assessed in quantitative (admixture) terms, rather than population estimates in dichotomous terms and rather than measures of "confounded" ancestry.

The present inventors were the first to assemble such a panel (Frudakis, 2007), the first to apply these panels for large-scale population surveys (Frudakis, 2007), the first to apply them in family pedigrees (Frudakis, 2007), and the first to relate measures of ancestry to anthropometric phenotypes demonstrating that, while measures of genomic ancestry were only correlated with anthropometric phenotypes, the expression of the gene variants underlying variable expression of the phenotypes are inextricably linked (see, Frudakis et al., Hum Genet (2007)122(3-4):311-326). Without the use of large panels of neutral markers, and highly precise measures of ancestry within individuals (in a quantitative sense) this formal proof would not have been possible to obtain. Not only were Applicants the first to measure true ancestry within individuals, Applicants were the first to prove measurement of true ancestry within individuals.

These demonstrations were not trivial, nor were they obvious. The existence of confounded population structure does not necessarily anticipate the existence of population structure due to genomic ancestry – since it could have been the result that, unlike other animals, population structure in humans is due only to differences of genes and in phenotypes. Applicants might even have expected this since humans are known from the paleoarchaeological record to share a very recent common ancestor compared to any other species of animal (Frudakis, 2007; Jobling et al., 2004). The first peer-reviewed article employing large-scale population surveys of admixture using neutral, autosomal markers appeared in 2003, written by Noah Rosenberg and colleagues at Stanford. This paper, which used microsatellites and was published after the

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page   5

PATENT
Attorney Docket No.: DNA1170-2

present application was filed, was and still is considered a landmark publication for the study of human evolution using individual ancestry admixture estimates (Rosenberg et al., Science (2002) 20:298(5602):2381-2385). Applicants work with neutral loci is equally important, and carried out before this work by Rosenberg, but contributed to the present patent application rather than to publications; indeed the presently discussed work was not submitted to journals until well after the present patent application was published (Halder et al., 2007 in press at Human Mutation; Frudakis 2007).

The findings by Rosenberg et al. (2003) were considered to be such an achievement the paper appeared in the journal Science. Indeed the observations from that work, and Applicants' unpublished work which preceded it, demonstrated that properly uncoupled ancestry from phenotype and geography showed very interesting patterns of apportionment across the globe – patterns which were far more informative than previously suspected based on measures of "confounded" ancestry. The observations on ancestral relationships among the world's inhabitants, as opposed to gene or phenotype relationships, provided a unique view of our history as a species not previously available. It was the novelty of this view, and the fact that Applicants had assembled and used such data prior to the Rosenberg paper, that induced Academic Press/Elsiever publishers to invite one of the Applicants to author the first textbook on the subject of resolving and inferring individual ancestry admixture and phenotype (Frudakis, 2007). Still, to this day, the work in this book (derived strictly from the invention subject of the patent application discussed herein) is the first to elucidate the relationships between genomic ancestry within individuals, geography, and social constructs such as self-assessed "ethnicity" or other geopolitical affiliations – wherefrom it is apparent that the patterns of apportionment of ancestry across "ethnic" and "racial" lines is even more interesting (Frudakis, 2007).

Another indication that Applicants' work is novel and unobvious is from the newsworthy nature of the application of the methods of the present invention for assisting with the resolution of several serial homicide cases. A "Google" search can illustrate several of these stories. For example, the methods and markers which are subjects of present application were instrumental in resolving the Louisiana Serial Killer Case, which was covered by CBS Evening News, ABC's

Prime Time Live, Forensic Files (Court TV program), New York Times (2 articles), Popular Science and in various international TV shows and newspapers. These stories covered the case as a historical first - billing the application of Applicants' marker panels and methods in the field of forensics as an example of the utility of novel inventions derived from the human genome project. Prior to this, measures of markers linked with phenotypically active loci among individuals were common as the Examiner's cited prior art attests, yet these measures had not been applied for forensics casework. The reason is that the authors of these manuscripts recognized that what they were measuring was a confounded variable, and the relationship between their confounded variable and phenotype would be impossible to formally define and thus, defend in a court of law.

The Examiner lists several manuscripts claimed to represent prior art, which make Applicants' claims unpatentable. As with any technological development, Applicants realize that the present accomplishments are built upon the benefit, in part, of work by those that came before. While these works made Applicants work possible, they do not make the present claims obvious. Applicants' findings required the investment of years of effort conducting large-scale population surveys, dissecting human phenotypes, building databases and conducing myriad validation assessments. Only from the fruits of this work can it be claimed that one is measuring genomic ancestry as opposed to "confounded" ancestry, which is not ancestry at all, notwithstanding mainly unsubstantiated claims by authors at the time these articles were published.

Applicants submit that the finding of SNPs, and the construction of SNP databases does not make the demonstration of an association between a given SNP and a disease, for example, obvious. It enables it, but does not instruct as to the existence of the association, nor specifically how it can be found. For example, Parra et al. (2001) is claimed to measure ancestry admixture but in fact, since the markers were from gene regions, some of them pigmentation genes, they measured "confounded" ancestry. The measurement of confounded ancestry neither makes the measurement of true ancestry obvious nor trivial because from the former, the latter is not necessarily apparent nor does it necessarily follow, as described above. Also as described,

detection of population structure due solely to genomic ancestry requires the use of neutral

AIMs, and is validated as such using large-scale global population and pedigree studies, with

reference to carefully quantified anthropometric phenotypes.  Parra et al. (2001) claim to teach

how Ancestry Informative Markers might be selected, but neglect to consider that such markers

should be neutral, found outside of gene regions and as such, they do not formally teach any such

thing.  Rather, they teach how markers for coalescing gene histories and reconstruction of

phenotype evolution can be found.  That is, rather than teaching how "AIMs" could be used to

infer ancestry, this paper teaches how gene markers that differ in frequency among populations

of disparate phenotypes might be used to infer aspects of phenotype evolution and/or the

evolution of shared environments.   Thus, the Examiners statement that "Parra et al., 2001

discloses a method for inferring ancestral proportions and admixture in six different populations

from different regions" is not correct (notwithstanding the claims to the contrary in Parra et al.

(2001)).   Parra et al. (2001) do not teach that population structure due to genomic ancestry is

distinct from population structure caused by shared environments and/or phenotypes.  Their

method of using delta > 0.4 is distinct from Applicants, because it is applied to gene sequences

rather than neutral mutations, and so the delta measures a different parameter (gene sequence and

phenotype difference, not ancestry difference).  Parent samples are used to estimate gene

sequence frequencies, but not AIM allele frequencies (notwithstanding claims to the contrary in

the paper).  The use of unlinked markers within genes to infer the evolution of gene sequences

and phenotypes does not make obvious the use of unlinked AIMs across the genome for the

inference of ancestry.  African American populations showing similarity to one another in terms

of phenotype and pigment gene sequences is expected, as is it expected they would show

similarity to one another in terms of ancestry.  This paper does not resolve the two, nor

demonstrate that the latter even exists.  The confidence intervals discussed in Parra et al. (2001),

and those after it preceding Applicants' application are fundamentally defective and incorrect.

Admixture is a multidimensional variable, and so too are the confidence intervals.  As such, they

must be considered in multidimensional space – for example for 3 intervals A:1-5, B:2-6 and

C:4-8, it is not the case that A:1 is equally likely to be seen with B:2 as with B:5 or C:6.  A:1,

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page 8

PATENT
Attorney Docket No.: DNA1170-2

B:2, C:4 may be within a 95% confidence interval, but A:1, B:2, C:5 may not be. Presenting the confidence intervals as A:1-5, B:2-6 and C:4-8 is therefore incorrect because it implies that all combinations incorporating values within each of the 3 ranges is equally likely. Admixture confidence intervals are best presented in terms of confidence contours in multidimensional space, such as that used in Applicants' triangle or tetrahedron plots (which can assume 3-dimensional and even 4-dimensional spaces on a 2-D piece of paper as described in Frudakis, 2007). The claim by the Examiner that "Parra et al., 2001 discloses biogeographical ancestry trait" as in claim 90 is incorrect, since gene sequence markers are used. For the same reason, the statement by the Examiner that "(Parra et al., 2001 discloses) proportional ancestry comprising a three-way comparison of sub-populations of African-Americans and the distribution percentage of European (ancestry) alleles within this sub-population derived from maximum-likelihood methods" is also unsubstantiated and incorrect.

The methods of SNP selection based on delta values and the likelihood algorithm used by Parra et al. (2001) had been developed by others before them (as far back as 1931, by Bernstein, *Die geographische Verteilung der blutgruppen und ihre anthropologische bedeutung. In: Comitato italiano per lo studio dei problemi della populazione*, 1931, Roma: Instituto Poligrafico dello Stato. 227–243), and happen to be useful for the purpose of inferring true ancestry if properly applied, as Applicants have applied them (i.e., in modified form, calculating the multidimensional confidence intervals or contours). However, they do not teach how legitimate Ancestry Informative Markers (as opposed to Phenotype Informative) are to be identified, since doing so requires focus on the neutral portion of the human genome.

The references cited by the Examiner do not make Applicants' invention obvious, because they do not recognize the difference between "confounded" and true ancestry, focus on the study of gene sequence and phenotype evolution rather than ancestry, do not measure ancestry, and do not teach how ancestry can be measured. The existence of phenotype evolution does not anticipate or make obvious the existence of genetic structure due to ancestry because phenotypes are acted upon by environment and their evolution uncoupled from ancestry. For example, phenotypes are well known to vary within species and even within regional clades

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page   9

PATENT
Attorney Docket No.: DNA1170-2

within species groups if there is corresponding environmental variation. The realization that confounded and true ancestry are not one and the same, and that the latter existed at all, required extensive R&D investment – the identification of neutral alleles, their application in world-wide population surveys, and their relationship with carefully assessed anthropometric  phenotypes (as well as other validation exercises).  At the time of the present application, none of these surveys or assessments had been executed by others and in fact, the difference between the two had not been formally demonstrated.  It was because such demonstration was not yet in hand that debate raged at the time over the very existence of bona-fide phylogenetic (e.g. genetic) structure within the human population, as opposed to variation at a few phenotype-relevant loci (which was obvious), or variation among human sex-chromosomes.  The latter was art recognized, but did not influence the debate because the measures do not apply to individuals (sex chromosomes are one of 23 chromosomes contributed by our ancestors, and so ancestry estimates for individuals are not possible), and because our species is highly sexually dimorphic, and gender related genes/chromosomes are subject to strong selective forces which do not necessarily apply to other chromosomes (indeed, phylogeny inferred from Y and mtDNA chromosomes are highly discordant from one another –Boissinot and Boursot 1997).

Further, while Pritchard (2000) and colleagues are the pioneers of the software systems most used for this purpose today, Applicants' statistical methodology differs from Pritchard (2000) in that the present confidence intervals for estimates are presented in multidimensional space.  Applicants were the first to do this, and this is the only correct way to do it as explained above.

Applicants claim to be the first to discriminate human population structure within individuals or populations of individuals, due to genetic ancestry as opposed to confounded structure due to shared phenotypes, gene sequences and environments.  The Examiner claims that Pritchard's (2000) results carried out in birds, in part, make Applicants' work in humans obvious and the observed results "predictable."  Pritchard's (2000) results do not make the present results predictable, because, as has been explained already, the existence of human population structure due to human genetic ancestry was not known to exist from Pritchard's (2000) work, and could

not be inferred to exist based on Pritchard's (2000) results from birds (or for any other species of animal). That it did exist was not confirmed until Applicants performed methods as disclosed in the instant application, separating true ancestry from confounded population structure in large scale human population studies. As discussed, prior to present application and the landmark study of Rosenberg et al. (2003), there was considerable debate as to the very existence of human population structure, since preceding work was confounded by the use of markers located within gene regions (i.e., not neutral) and since humans were/are believed to have derived from common ancestors much more recently than other species of animals (e.g., such as birds). Debates raging on the topic at the time and even after (see, e.g., Cho and Sankar, 2004; Shriver et al., 2005) indicate this uncertainty and establish that the results of Rosenberg et al. (2003) were the first to address it in the peer-reviewed literature. Pritchard's (2000) results do not anticipate or make obvious Rosenberg's in 2003, nor the Applicants', or even teach how or whether Applicants' or those of Rosenberg et al. (2003) would/could obtain such results, since Pritchard et al. (2000) did not work on human populations, and since the human population is relatively recently derived from common ancestors compared with other species, it is the most difficult within which to resolve structure (particularly phylogenetic structure, created by diversity with respect to true, genetic ancestry) Applicants' and Rosenberg's (2003) results required identification of true AIMs (i.e., markers located outside gene regions) and the study of global populations, and the validation that the elements of structure observed did not strictly comport with phenotype, geography or social variables. Using evolutionary genetics results from species such as birds to form predictions about human populations is not a standard protocol within the evolutionary genetics community. As with comparing apples and oranges, using one species with a unique demographic and evolutionary history to make inferences about another is scientifically untenable, and could not be published in any respectable peer-reviewed journal.

For something to be obvious, it has to be predictable and Applicants have argued in this rebuttal that in the instant case, the prior results measuring confounded population structure in humans or population structure of indeterminate origin in other species such as birds do not predict Applicants' results measuring human population structure due to genomic ancestry.

In re Application of
    Frudakis and Shriver
Application No.: 10/644,594
Filing Date: August 19, 2003
Page  11

PATENT
Attorney Docket No.: DNA1170-2

Pritchard's other cited paper in Theoretical Population Biology (2001) does not appreciate nor address the distinction between confounded population structure and structure due to ancestry.  This review discusses the statistical theory of how measures of population structure could positively impact case-control study designs, but in relying on the body of work preceding it, as has been discussed, is focused on measures of confounded ancestry and shared environment, the paper is more about correcting for population structure created by shared environment or convergent evolution than structure created by phylogenetic diversity, which is what we measure using neutral markers and an assessment of genetic ancestry.  The methods may or may not be applied to both types of structure, but the discussion of the former or even the latter does not make obvious the existence of the latter, as discussed.  Further, the paper was published after Applicants' patent application priority date (November 2001 vs. May, 2001).

5.      I further declare that all statements made herein of knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine, or imprisonment, or both under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

_____1/9/08_____
Date

_____7-7.·_____
Tony Frudakis, Ph.D.